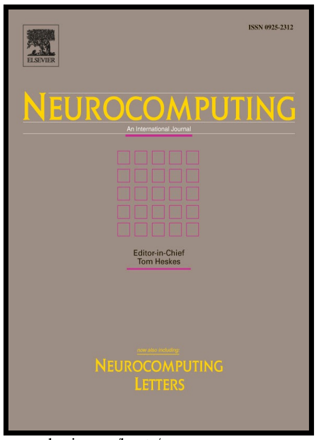
Author's Accepted Manuscript

A Two-Stage Model for Time Series Prediction based on Fuzzy Cognitive Maps and Neural Networks

Elpiniki I. Papageorgiou, Katarzyna Poczęta



www.elsevier.com/locate/neucom

PII: S0925-2312(16)31575-2
DOI: http://dx.doi.org/10.1016/j.neucom.2016.10.072
Reference: NEUCOM17878

To appear in: *Neurocomputing*

Received date: 30 April 2016
Revised date: 11 October 2016
Accepted date: 19 October 2016

Cite this article as: Elpiniki I. Papageorgiou and Katarzyna Poczęta, A Two-Stage Model for Time Series Prediction based on Fuzzy Cognitive Maps anc Neural Networks, *Neurocomputing* http://dx.doi.org/10.1016/j.neucom.2016.10.072

A Two-Stage Model for Time Series Prediction based on Fuzzy Cognitive Maps and Neural Networks

Elpiniki I. Papageorgiou[a,b,*], Katarzyna Poczęta[c]

[a] *Technological Educational Institute (T.E.I.) of Central Greece, 3rd Km Old National Road Lamia-Athens, Lamia 35100, Greece*
[b] *Hasselt University, Campus Diepenbeek Agoralaan Gebouw D, BE3590 Diepenbeek, Belgium*
[c] *Kielce University of Technology, al. Tysiąclecia Państwa Polskiego 7, 25-314 Kielce, Poland*

Abstract

This paper proposes a two-stage prediction model, for multivariate time series prediction based on the efficient capabilities of evolutionary fuzzy cognitive maps (FCMs) enhanced by structure optimization algorithms and artificial neural networks (ANNs). In the first-stage, an evolutionary FCM is constructed automatically from historical time series data using the previously proposed structure optimization genetic algorithm, while in the second stage, the produced FCM defines the inputs in an ANN which next is trained by the back propagation method with momentum and Levenberg-Marquardt algorithm on the basis of available data. The structure optimization genetic algorithm for automatic construction of FCM is implemented for modeling complexity based on historical time series data, selecting the most important nodes (attributes) and interconnections among them thus providing a less complex and efficient FCM-based model. This models is used next as input in an ANN. ANNs are used at the final process for making time series prediction considering as inputs the concepts defined by the produced FCM. The previously proposed structure optimization genetic algorithm for FCM construction by historical data as well as the ANN have been already proved their efficacy on time series forecasting.

*Corresponding author
Email address: epapageorgiou@teiste.gr (Katarzyna Poczęta)

The performance of the proposed approach is presented through the analysis of multivariate historical data of benchmark datasets for making predictions. The multivariate analysis of historical data is held for a large number of input variables, like season, month, day or week, holiday, mean and high temperature, etc. The whole approach was implemented in an intelligent software tool initially deployed for FCM prediction. Through the experimental analysis, the usefulness of the new two-stage approach in time series prediction is demonstrated, by calculating seven prediction performance indicators which are well known from the literature.

Keywords: fuzzy cognitive map, artificial neural network, forecasting, time series prediction, real coded genetic algorithm

---

1. Introduction

In recent years, different methods and approaches to multivariate time series forecasting have been proposed. These methods vary from the simplest historical extrapolation to sophisticated analytical models; the choice of an appropriate model depends on the purpose of the forecasting required by the investigated problem, as well as the quality and quantity of data. Recently, forecasting methods have been developed mainly using autoregressive integrated moving average (ARIMA) model, artificial neural network (ANN), neuro-fuzzy systems and other hybrid approaches, based on time-series data that are collected sequentially over various time periods [50].

Much effort has been devoted to develop and improve the hybrid time series models, which usually combine the linear and nonlinear predictors in a cascade form in order to increase time series prediction accuracy. Most of the hybrid models or two-stage models are combining two or more of the well-known forecasting methods, like ARIMA, ANNs and regression models [4, 17, 18, 19, 47]. For example, in [24] the hybrid model considers the routine time prediction technique like autoregressive model (AR), ANN or any others as atomic building block. Wang *et al.* (2013) proposed a hybrid model, which is distinctive in integrating the advantages of ARIMA and ANNs in modeling the linear and nonlinear behaviors in economic data set [46]. The proposed approaches consider the linear and nonlinear patterns simultaneously in real data, so that they can mine more precise characteristics to describe the time series modeling better.

Recently, evolutionary Fuzzy Cognitive Maps have been investigated to model time series problems and make forecast. FCMs work as recurrent neural networks, inheriting their main advantages through the learning capabilities. They model any real world system as a collection of concepts and causal relationships among concepts [20]. They were initially relied on the human expert knowledge for a domain, making associations along generalized relationships between domain descriptors, concepts and conclusions [35, 37, 38, 44, 45, 48]. However, due to the availability of historical data and supervised and population-based algorithms for modifying the relationships between the concepts, intelligent methodologies have been exploited for automatic construction of FCMs (Papageorgiou *et al.*) [25, 26, 27, 28]. Their learning capabilities enhanced the operation of FCMs for modeling and prediction tasks [39, 42] and helped them to gain momentum due the last years [31].

The application of FCMs to time series modeling was initially discussed by Homenda *et al.* [13, 14, 21]. In [14], nodes selection criteria for FCM designed to model univariate time series were proposed. Also some simplifications strategies by posteriori removing nodes and weights were presented in [14]. Salmeron and Froelich proposed an FCM-based approach for the forecasting of univariate time series with the dynamic optimization of the FCM structure [36]. A hybrid algorithm for fuzzy time series prediction based on fuzzy c-means clustering, FCM and genetic algorithm was presented in [22]. Evolutionary FCM was also applied for the effective prediction of multivariate interval-valued time series [8]. In that work, an evolutionary algorithm for learning fuzzy grey cognitive maps (FGCMs) was developed as a nonlinear predictive model. The new algorithm was applied to learn FGCMs on the basis of meteorological time series data providing evidence that, for properly-adjusted learning and prediction horizons, the proposed approach can be used effectively to the forecasting of multivariate, interval-valued time series.

Moreover, alternative approaches to FCM-based time series modeling are related to classification. It deserves to pinpoint that the first research paper on FCM-ANN, as a hybrid model, was proposed by Papakostas *et al.* for pattern classification [32]. The results provided by the conducted experiments in pattern recognition showed that this type of hybrid technique allowed to improve the operation of the model based on both FCM and ANN capabilities.

Recently, evolutionary FCMs enhanced by Structure Optimization Genetic Algorithm (SOGA) have been proposed in [29, 34]. SOGA enables fully automatic construction of the FCM model by selection of crucial concepts and determination of the relationships between them on the basis of available historical data. SOGA-FCM was compared with some well-known methods for FCM construction from data: the Multi-Step Gradient Method (MGM) [16] and the Real-Coded Genetic Algorithm (RCGA) [40] on the example concerning the prediction of count of rented bikes.

The investigated study is focused on the use of evolutionary FCM and ANNs in a subsequent process for multivariate time series prediction. Based on the literature, ANNs and FCM have proved their efficiency in prediction; however, no research has been done on combining the advantageous characteristics and dynamic capabilities of each one to a new two-stage process. The two-stage structure consists at the first stage of a SOGA-FCM approach for constructing FCM by historical data using population-based method and at the second stage from an ANN which inputs are defined by the first stage model.

The aim of this research study is to simplify the problem of multi-variate time series prediction by removing the redundant concepts (data attributes), proposing a less complex model consisting of the most important attributes defined by historical time series data, which will be used as input in an ANN for providing higher accuracy on prediction for multi-variate time series data.

To succeed the goal of this study, a two-stage process based on evolutionary FCM and ANN was proposed inheriting the main advantageous characteristics of FCMs and ANNs, with efficient learning algorithms based on gradient-based methods and population-based methods. Through the two-stage method the most significant attributes are selected by the evolytionary construction of FCMs based on SOGA which are used next as inputs in an ANN model trained by an efficient algorithm. At first, an improved version of the SOGA for FCM learning is used to select the most important input data (concepts) which produce acceptable accuracy on prediction (through well-known performance indicators). SOGA-FCM allows to select the most significant for the modeled system concepts (attributes) and connections between them. The effectiveness of this method was presented on a previously research study concerning time series prediction [29, 34]. The role of the produced FCM after SOGA learning is to simplify the problem and remove the redundant concepts (data attributes). Thus, the proposed SOGA-FCM provides the selected concepts to the ANN as inputs.

Four multivariate time series, historical datasets, well-known from the literature consist of a relatively large number of concepts that need to be assessed and analyzed for multi-variate time series prediction, were considered for evaluation of the new proposed approach. The first dataset concerns the prediction of the electric power consumption (reported in [11]), the second dataset concerns the prediction of stock exchange index [2, 3], the third is devoted to the time series prediction of the bike sharing counts [5] and the last one refers to the indoor temperature forecasting towards energy efficiency [49].

A comparative analysis of the two-stage method with the well-known ANN and FCM learning algorithms (Real-Coded Genetic Algorithm and SOGA-FCM) was performed with the use of ISEMK (Intelligent Expert System based on Cognitive Maps) software tool. Through the experimental analysis, the usefulness of the two-stage approach in time series prediction is demonstrated, by calculating seven prediction performance indicators which are well known from the literature. The obtained results show that the proposed approach allows to significantly reduce the problem complexity through the designed structure of the FCM model by selecting the most important concepts, which are used in ANN to make the final forecast. The results of the conducted experiments provide a relatively high forecasting accuracy, comparing with the accuracy provided by ANN or FCM as indepented forecasting approaches.

The outline of this paper is as follows. Section II briefly describes fuzzy cognitive maps and evolutionary learning algorithms. Section III presents the learning algorithms for artificial neural networks. In Section IV, we introduce the two-stage approach for time series prediction based on FCMs and ANNs. Section V shows selected results of simulation analysis of the proposed method on four prediction problems reported in the literature. Section VI discusses the results and last section concludes the general remarks, including the directions of further research.

2. Learning Algorithms for Fuzzy Cognitive Maps

Fuzzy cognitive map is a directed graph described by the set of the concepts X and the connections matrix W [20]:

$$< X, W > , \tag{1}$$

where $X = [X_1, ..., X_n]^T$ is the vector with the values of the concepts, W is the connection matrix representing the relationships between concepts.

The concepts influence each other with an intensity described by the weight of the connection between them. Values of the concepts at the next iteration can be calculated on the basis of the current state vector (current values of the concepts) according to the selected dynamic model. In the analysis, two popular models (2), (3) were used [31].

$$X_i(t+1) = F\left(X_i(t) + \sum_{j \neq i} w_{j,i} \cdot X_j(t)\right) , \tag{2}$$

$$X_i(t+1) = F\left(\sum_{j \neq i} w_{j,i} \cdot X_j(t)\right) , \tag{3}$$

where $w_{j,i}$ is the weight of the connection between the j-th concept and the i-th concept, taking on the values from the range $[-1, 1]$, $X_i(t)$ is the value of the i-th concept at the t-th iteration, t is discrete time, $t = 0, 1, 2, ..., T$, T is end time of simulation, $i = 1, 2, ..., n$, n is the number of concepts, $F(x)$ is a transformation function, which normalizes the values of the concepts to the range [0,1]. It can be chosen in the form:

$$F(x) = \frac{1}{1 + e^{-cx}} , \tag{4}$$

where $c > 0$ is a parameter.

One of the most important advantages of the fuzzy cognitive maps is the ability for automatic construction of the model based on historical data. It can be achieved by supervised [16, 33, 43] and evolutionary learning algorithms [1, 7, 8, 40].

In the paper, the popular Real-Coded Genetic Algorithm and the developed Structure Optimization Genetic Algorithm for FCM learning were analyzed.

2.1. Real-Coded Genetic Algorithm (RCGA)

Real-Coded Genetic Algorithm was used by Stach *et al.* in 2005 to train the FCM model [40]. It allows to determine the matrix connections for the given set of concepts based on the available data. The aim of the learning process is to compute the FCM model that is able to mimic the input data.

RCGA defines each individual as a floating-point vector, determined based on connection matrix [40]:

$$W' = [w_{1,2}, ..., w_{1,n}, w_{2,1}, w_{2,3}, ..., w_{2,n}, ..., w_{n,n-1}]^T , \tag{5}$$

where $w_{j,i}$ is the weight of the connection between the j-th and the i-th concept.

In this paper, additional parameters of every individual were analyzed:

- c – parameter of the transformation function (4),
- d_m – the dynamic model, $d_m = 0$ means that the model type (2) is set, $d_m = 1$ means that the model (3) is set.

Each individual is decoded into a candidate FCM and evaluated based on a fitness function [40]. This function is designed with taking into account the objective of the research [6].

The following fitness function was proposed in [40]:

$$fitness_p(J(l)) = \frac{1}{a \cdot J(l) + 1} , \tag{6}$$

where a is a parameter established experimentally, l is the number of generation, $l = 1, ..., L$, L is the maximum number of generations, p is the number of individual, $p = 1, ..., P$, P is the population size, $J(l)$ is the error measure.

The core of the fitness function is error measure $J(l)$. The aim of the research is the one-step-ahead prediction, so the error function is based on the difference between the candidate FCM response (values of the output concepts) and historical normalized data (desired values of the output concepts), described as follows:

$$J(l) = \frac{1}{(T-1)n_o} \sum_{t=1}^{T-1} \sum_{i=1}^{n_o} (Z_i^o(t) - X_i^o(t))^2 , \tag{7}$$

where t is discrete time of learning, T is the number of the learning records, $Z(t) = [Z_1(t), ..., Z_n(t)]^T$ is the desired FCM response for the initial vector $Z(t-1)$, $X(t) = [X_1(t), ..., X_n(t)]^T$ is the FCM response for the initial vector $Z(t-1)$, n is the number of the concepts, n_o is the number of the output concepts, $X_i^o(t)$ is the value of the i-th output concept, $Z_i^o(t)$ is the reference value of the i-th output concept.

The RCGA stops when the learning is successful which means that the best candidate FCM response is satisfactory (close to the historical time series) or the maximum number of generations L is reached.

2.2. Structure Optimization Genetic Algorithm (SOGA)

The structure optimization genetic algorithm is an extension of the RCGA method for FCMs learning [29, 34]. Real-coded genetic algorithm allows to determine the connection matrix based on the available data. It requires identifying the concepts of the map by experts or the FCM model is created on the basis of all available data attributes. However, FCMs with large number of concepts and connections between them can be unreadable and difficult to be interpreted [14]. The developed SOGA algorithm allows to select only the most significant concepts (data attributes) and connections between them keeping similar level of error measures. The resulting FCMs are less complex regarding the number of concepts and connections [34].

SOGA defines each individual using the floating-point vector type (5), parameters c, d_m and a binary vector expressed by the formula (8). This binary vector allows to select random concepts of all possible vectors based on a specified probability P_{zero}. This probability determines the possible number of concepts in the FCM model.

$$C = [C_1, C_2, ..., C_n]^T \ , \tag{8}$$

where C_i is the information about including the i-th concept to the candidate FCM model, whereas $C_i = 1$ means that the candidate FCM model contains the i-th concept, $C_i = 0$ means that the candidate FCM model does not contain the i-th concept (the corresponding weights of the connections are also equal 0). For the output concepts, the values are equal to 1.

Each individual is decoded into a candidate FCM that contains concepts indicated by vector C. The quality of every individual is calculated based on the fitness function, described as follows:

$$fitness_p(J'(l)) = \frac{1}{a \cdot J'(l) + 1} \ , \tag{9}$$

where a is a parameter established experimentally, l is the number of generation, $l = 1, ..., L$, L is the maximum number of generations, p is the number of the individual, $p = 1, ..., P$, P is the population size, $J'(l)$ is the new error measure.

This error function takes into account an additional penalty for highly complexity of FCM understood as a large number of concepts (n_c) and non-zero connections between them (n_r) [34]:

$$J'(l) = J(l) + b_1 \cdot \frac{n_r}{n^2} \cdot J(l) + b_2 \cdot \frac{n_c}{n} \cdot J(l) \ , \tag{10}$$

where b_1, b_1 are the parameters established experimentally, $b_1 > 0$, $b_2 > 0$, n_r is the number of the non-zero connections, n_c is the number of the concepts in the candidate FCM model, n is the number of all concepts, $J(l)$ is the learning error function type (7).

Using a binary vector C and the proposed error function the redundant concepts and connections between them are removed and less complex model for time series prediction is constructed.

The construction of the FCM model with the use of genetic algorithms consists of the following steps:

- initialize the FCM model based on all available historical time series,
- determine learning parameters,
- initialize population,
- generate new population using evolutionary operators and selection strategy,
- evaluate population based on the developed fitness function,
- check stop condition.

The following evolutionary operators were used in our experiments [6, 12]:

- ranking selection,
- uniform crossover,
- random mutation.

Additionally, elite strategy of selection was applied. Each population was assigned a probability of crossover P_c and mutation P_m.

3. Learning Algorithms for Artificial Neural Networks

Artificial neural networks are data-driven models that can be learned with the use of supervised learning algorithms and historical data. The two most popular supervised learning techniques for ANNs for prediction are: back propagation method with momentum [10] and Levenberg-Marquardt algorithm [9].

The basis of the structure of ANN is a neuron, that is an information-processing unit. It contains a set of synapses (connecting links). The input signal x_j connected to neuron k is multiplied by the synaptic weights w_{kj}. The output of the neuron y_k can be described as follows [10]:

$$y_k = F\left(\sum_{j=1}^{m} w_{kj} \cdot x_j + b_k\right) , \tag{11}$$

where w_{kj} is the synaptic weight, x_i is the input signals, $i = 1, 2, ..., m$, m is the number of signals, b_k is the bias, $F(x)$ is an activation function, that can be chosen in the form (4).

In a layered neural network the neurons are organized in the form of layers. The most widely used ANNs in time series prediction are multilayer perceptrons with an input layer, a single hidden layer and an output layer. The input signals propagate through the network in a forward direction, on a layer-by-layer basis.

3.1. Back Propagation Method with Momentum

One of the most popular methods for ANNs learning is back propagation algorithm [10]. It is a method of learning in which the weights are updated based on the learning records until one epoch (one complete presentation of entire learning dataset has been deal with). The objective of the learning process is to minimize the learning error function described as follows:

$$J(l) = \frac{1}{2 \cdot (T-1)} \sum_{t=1}^{T-1} \sum_{k=1}^{n_o} (Z_k(t) - y_k(t))^2 , \tag{12}$$

where t is discrete time of learning, T is the number of the learning records, $Z_i(t)]^T$ is the desired ANN response, $y_k(t)$ is the ANN response, n_o is the number of the output neurons, l is the number of epoch, L is the maximum number of epoch.

In back propagation algorithm, the modification of the weights is based on the formula:

$$\Delta w_{kj}(l) = -\gamma \cdot \frac{\partial J(l)}{\partial w_{kj}(l)} , \tag{13}$$

where $\Delta w_{kj}(l)$ is a change of the weight w_{jk} at the lth epoch, γ is a learning coefficient.

Back propagation algorithm with momentum modifies the weights according to the formula:

$$\Delta w_{kj}(l) = -\gamma \cdot \frac{\partial J(l)}{\partial w_{kj}(l)} + \alpha \Delta w_{kj}(l-1) \ , \tag{14}$$

where α is a momentum parameter.

3.2. Levenberg-Marquardt Algorithm

Levenber-Marquardt algorithm combines the back-propagation algorithm with the Newton method. It modifies the weights on the basis of the following formula [9]:

$$w_{l+1} = w_l - (J_l^T J_l + \mu_l I)^{-1} J_l^T E_l \ , \tag{15}$$

where J is Jacobian of the k output errors with respect to n weights of ANN, E is the cumulative error vector, μ is a learning parameter. For $\mu = 0$ the process becomes the Newton method. For large μ the algorithm works as back propagation method.

The supervised learning stops when the process is successful (the learning error J is low) or the maximum number of epochs L is reached.

4. Two-Stage Approach for Time Series Prediction based on SOGA-FCM and ANN

This section presents the proposed two-stage process for time series prediction. The aim of this approach is:

- to remove the redundant concepts and construct readable and clear model for time series prediction based on SOGA-FCM,
- to improve the prediction accuracy of the ANN model by using inputs from SOGA-FCM,
- to find the most accurate model of two stage process for time series prediction problems.

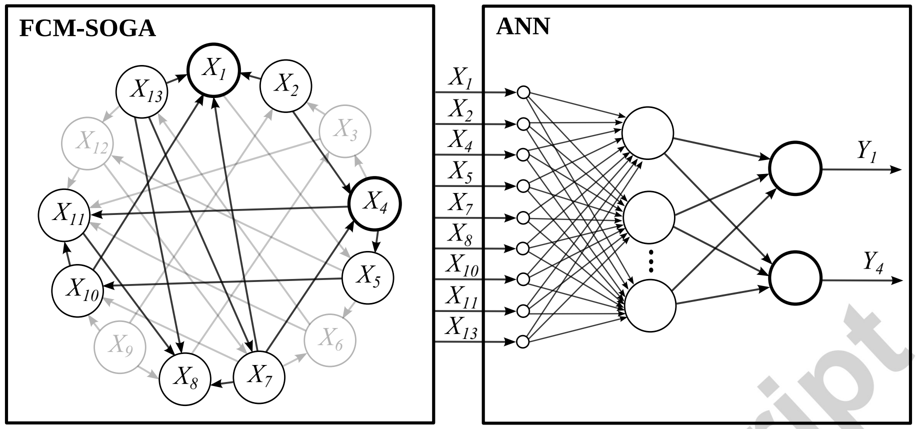


Figure 1: Two-stage process for multivariate time series prediction

Figure 1 illustrates the proposed approach for multivariate time series prediction.

The developed algorithm consists of two stages described below.

STAGE 1. *Construct the FCM model automatically from historical time series using the structure optimization genetic algorithm.*

The first stage is a process described as follows:

1. Initialize the FCM model based on all available historical time series.
2. Build the optimized FCM model (select the most significant concepts and connections between them) using the SOGA algorithm.
3. Test the resulted FCM with the use of normalized testing data.

STAGE 2. *Use the concepts of the produced FCM model to determine the inputs for ANN in order to increase the prediction accuracy.*

The second stage is a process described as follows:

1. Increase the prediction accuracy by using the selected input attributes (concepts of the resulted FCM model) as an input dataset for artificial neural network.
2. Learn ANNs with the use of back propagation method with momentum [10] or Levenberg-Marquardt algorithm [9] on the basis of selected by SOGA historical time series.

3. Test the resulted ANN with the use of normalized testing data.

The experiments were carried out with the use of the developed software tool ISEMK in order to evaluate the quality of the proposed approach [15].

5. Selected Results

For the evaluation process, historical time series normalized with the use of min-max normalization 16 were used.

$$f(x) = \frac{x - min}{max - min} , \tag{16}$$

where x is an input numeric value, min is the minimum of the dataset, max is the maximum of the dataset.

5.1. Datasets

Four datasets with various number of concepts (attributes) were used to show and analyze the usefulness of the two-stage model. Tables 1 and 2 present the attributes of the analyzed data. The first dataset (Electricity) contains measurements of different electrical quantities and some sub-metering values for electric power consumption [11]. In the analysis, energy sub-metering No. 1 (X_6), energy sub-metering No. 2 (X_7) and energy sub-metering No. 3 (X_8) were chosen as the output concepts. The second dataset (Stock) includes returns of Istanbul Stock Exchange with eight other international indexes from January 5, 2009 to February 22, 2011 [2, 3]. The first concept X_1 was chosen as the output concept. The third dataset (Bikes) contains bike sharing counts aggregated on daily basis (731 days) [5]. We set count of users (X_{12}), count of registered users (X_{13}) and count of total rented bikes (X_{14}) as the output concepts. The core data set is related to the two-year historical log corresponding to years 2011 and 2012 from Capital Bikeshare system, Washington D.C., USA (http://capitalbikeshare.com/system-data). The corresponding weather and seasonal information (http://www.freemeteo.com) were added. The fourth dataset is collected from a monitor system mounted in a domotic house [49]. It corresponds to approximately 40 days of monitoring data and has the following output concepts: indoor temperature in dinning-room (X_5) and in room (X_6).

Table 1: Attributes of the datasets 1-3

Dataset	Concept	Description
1 Electricity	X_1	time (in UTC)
	X_2	household global minute-averaged active power (in kilowatt)
	X_3	household global minute-averaged reactive power (in kilowatt)
	X_4	voltage: minute-averaged voltage (in volt)
	X_5	household global minute-averaged current intensity (in ampere)
	X_6	energy sub-metering No. 1 (in watt-hour of active energy)
	X_7	energy sub-metering No. 2 (in watt-hour of active energy)
	X_8	energy sub-metering No. 3 (in watt-hour of active energy)
2 Stock	X_1	ISE100 Istanbul stock exchange national 100 index
	X_2	USD ISE
	X_3	SP Standard & poor's 500 return index
	X_4	DAX Stock market return index of Germany
	X_5	FTSE Stock market return index of UK
	X_6	NIK Stock market return index of Japan
	X_7	BVSP Stock market return index of Brazil
	X_8	EU MSCI European index
	X_9	EM MSCI emerging markets index
3 Bikes	X_1	season (1:springer, 2:summer, 3:fall, 4:winter)
	X_2	year (0: 2011, 1:2012)
	X_3	month (1 to 12)
	X_4	holiday (1: yes, 0: no)
	X_5	weekday, day of the week (1: Monday, 7: Sunday)
	X_6	working day (1: yes. 0: no)
	X_7	weather situation (1:Clear, 2: Mist, 3: Light rain, 4: Heavy rain)
	X_8	temperature (normalized temperature in Celsius)
	X_9	feeling temperature (normalized feeling temperature in Celsius)
	X_{10}	humidity (normalized humidity)
	X_{11}	wind speed (normalized wind speed)
	X_{12}	casual (count of casual users)
	X_{13}	registered (count of registered users)
	X_{14}	count (count of total rented bikes including both casual and registered)

Table 2: Attributes of the dataset 4

Dataset	Concept	Description
	X_1	day (1 to 31)
	X_2	month (1 to 12)
	X_3	year, (2012)
	X_4	time, in UTC
	X_5	indoor temperature (dinning-room), in ^oC
	X_6	indoor temperature (room), in ^oC
	X_7	weather forecast temperature, in ^oC
	X_8	carbon dioxide in ppm (dinning room)
	X_9	carbon dioxide in ppm (room)
	X_{10}	relative humidity (dinning room), in %
	X_{11}	relative humidity (room), in %
	X_{12}	lighting (dinning room), in Lux
4	X_{13}	lighting (room), in Lux
Temperature	X_{14}	rain, the proportion of the last 15 minutes (a value in range [0,1])
	X_{15}	sun dusk
	X_{16}	wind, in m/s
	X_{17}	sun light in west facade, in Lux
	X_{18}	sun light in east facade, in Lux
	X_{19}	sun light in south facade, in Lux
	X_{20}	sun irradiance, in $\frac{W}{m^2}$
	X_{21}	enthalpic motor 1, 0 or 1 (on-off)
	X_{22}	enthalpic motor 2, 0 or 1 (on-off)
	X_{23}	enthalpic motor turbo, 0 or 1 (on-off)
	X_{24}	outdoor temperature, in ^oC
	X_{25}	outdoor relative humidity, in %
	X_{26}	day of the week (1: Monday, 7: Sunday)

5.2. Evaluation criteria

The performance of the proposed approach was analyzed with the use of the following criteria:

1. Mean Absolute Error:

$$MAE = \frac{1}{n_D \cdot T} \sum_{t=1}^{T} \sum_{i=1}^{n_D} |Z_i^D(t) - X_i^D(t)| \tag{17}$$

2. Mean Squared Error:

$$MSE = \frac{1}{n_D \cdot T} \sum_{t=1}^{T} \sum_{i=1}^{n_D} \left(Z_i^D(t) - X_i^D(t)\right)^2 , \tag{18}$$

3. Root Mean Squared Error:

$$RMSE = \sqrt{\frac{1}{n_D \cdot T} \sum_{t=1}^{T} \sum_{i=1}^{n_D} \left(Z_i^D(t) - X_i^D(t) \right)^2}, \tag{19}$$

4. Mean Absolute Percentage Error:

$$MAPE = \frac{1}{n_D \cdot T} \sum_{t=1}^{T} \sum_{i=1}^{n_D} \left| \frac{Z_i^D(t) - X_i^D(t)}{Z_i^D(t)} \right|, \tag{20}$$

5. Coefficient of Correlation:

$$R = \frac{1}{n_D} \sum_{i=1}^{n_D} \frac{T \cdot \sum_{t=1}^{T} Z_i^D(t) \cdot X_i^D(t) - \sum_{t=1}^{T} Z_i^D(t) \cdot \sum_{t=1}^{T} X_i^D(t)}{\sqrt{T \cdot \sum_{t=1}^{T} (Z_i^D(t))^2 - (\sum_{t=1}^{T} Z_i^D(t))^2} \cdot \sqrt{T \cdot \sum_{t=1}^{T} (X_i^D(t))^2 - (\sum_{t=1}^{T} X_i^D(t))^2}}, \tag{21}$$

6. Coefficient of determination R^2,

340 7. Coefficient of Efficiency [23]:

$$CE = 1 - \frac{1}{n_D} \sum_{i=1}^{n_D} \frac{\sum_{t=1}^{T} (Z_i^D(t) - X_i^D(t))^2}{\sum_{t=1}^{T} (Z_i^D(t) - \bar{Z}_i^D)^2}, \tag{22}$$

where $X_i^D(t)$ is the value of the ith output concept at iteration t for the candidate FCM, $Z_i^D(t)$ is the desired value of the ith output concept at iteration t, $t = 1, ..., T$, T is the number of testing records, $i = 1, ..., n_D$, n_D is the number of output concepts and $\bar{Z}_i^D$ is the mean of the ith output concept.

5.3. Results of evolutionary FCM learning

The first stage of the proposed approach involves the construction of the FCM model with the use of the SOGA algorithm. The results of comparative analysis of the SOGA algorithm and the RCGA algorithm are presented in Tables 3 and 4. For every configuration of the parameters (mutation probability, crossover probability, population size, maximum number of generations, fitness function parameters) learning process was performed 10 times and the average values of evolution criteria with standard deviation were calculated. Uniform crossover, random mutation, ranking selection and elite strategy were used in the simulations. Due to zero values, the calculation of $MAPE$ measurement was impossible for the first dataset.

Table 3: Chosen results of the analysis of the evolutionary FCMs (RCGA and SOGA) - datasets 1,2

Dataset	Method	Parameters	Criteria	Errors	n_c	n_r
		$L = 100$	MSE	0.022 ± 0.001		
		$P = 500$	$RMSE$	0.147 ± 0.003		
		$a = 1000$	MAE	0.066 ± 0.007		
1	RCGA	$b_1 = b_2 = 0$	$-$	$-$	8	40 ± 2
Electricity		$P_m = 0.05$	R	0.571 ± 0.075		
		$P_c = 0.5$	R^2	0.326 ± 0.044		
		$P_{zero} = 0$	CE	0.365 ± 0.060		
		$L = 100$	MSE	0.021 ± 0.001		
		$P = 500$	$RMSE$	0.146 ± 0.002		
		$a = 1000$	MAE	0.066 ± 0.007		
1	SOGA	$b_1 = b_2 = 0.01$	$-$	$-$	6 ± 1	15 ± 5
Electricity		$P_m = 0.05$	R	0.599 ± 0.032		
		$P_c = 0.5$	R^2	0.358 ± 0.042		
		$P_{zero} = 0.2$	CE	0.383 ± 0.032		
		$L = 100$	MSE	0.036 ± 0.027		
		$P = 500$	$RMSE$	0.183 ± 0.053		
		$a = 1000$	MAE	0.161 ± 0.054		
2	RCGA	$b_1 = b_2 = 0$	$MAPE$	0.404 ± 0.123	9	53 ± 4
Stock		$P_m = 0.05$	R	0.244 ± 0.497		
		$P_c = 0.5$	R^2	0.060 ± 0.027		
		$P_{zero} = 0$	CE	-2.498 ± 2.041		
		$L = 100$	MSE	0.026 ± 0.005		
		$P = 500$	$RMSE$	0.160 ± 0.018		
		$a = 1000$	MAE	0.137 ± 0.018		
2	SOGA	$b_1 = b_2 = 0.01$	$MAPE$	0.349 ± 0.048	5 ± 1	12 ± 5
Stock		$P_m = 0.05$	R	0.223 ± 0.060		
		$P_c = 0.5$	R^2	0.049 ± 0.027		
		$P_{zero} = 0.2$	CE	-1.498 ± 0.487		

The best FCM models (with the lowest value of MSE) were selected for further analysis. The most significant concepts defined by the SOGA-FCM approach were chosen to be further analyzed and depicted in Table 5.

Table 4: Chosen results of the analysis of the evolutionary FCMs (RCGA and SOGA) - datasets 3,4

Dataset	Method	Parameters	Criteria	Errors	n_c	n_r
		$L = 100$	MSE	0.040 ± 0.006		
		$P = 200$	$RMSE$	0.199 ± 0.015		
		$a = 10000$	MAE	0.153 ± 0.014		
3	RCGA	$b_1 = b_2 = 0$	$MAPE$	0.999 ± 0.206	14	124 ± 5
Bikes		$P_m = 0.05$	R	0.537 ± 0.074		
		$P_c = 0.5$	R^2	0.288 ± 0.097		
		$P_{zero} = 0$	CE	0.197 ± 0.045		
		$L = 100$	MSE	0.044 ± 0.005		
		$P = 200$	$RMSE$	0.209 ± 0.013		
		$a = 10000$	MAE	0.163 ± 0.013		
3	SOGA	$b_1 = b_2 = 0.01$	$MAPE$	0.956 ± 0.183	12 ± 1	79 ± 16
Bikes		$P_m = 0.05$	R	0.486 ± 0.089		
		$P_c = 0.5$	R^2	0.236 ± 0.076		
		$P_{zero} = 0.2$	CE	0.131 ± 0.100		
		$L = 100$	MSE	0.018 ± 0.003		
		$P = 100$	$RMSE$	0.132 ± 0.012		
		$a = 1000$	MAE	0.106 ± 0.010		
4	RCGA	$b_1 = b_2 = 0$	$MAPE$	0.56 ± 0.086	26	427 ± 16
Temperature		$P_m = 0.05$	R	0.863 ± 0.021		
		$P_c = 0.5$	R^2	0.744 ± 0.037		
		$P_{zero} = 0$	CE	0.616 ± 0.088		
		$L = 100$	MSE	0.013 ± 0.002		
		$P = 100$	$RMSE$	0.115 ± 0.01		
		$a = 1000$	MAE	0.093 ± 0.008		
4	SOGA	$b_1 = b_2 = 0.01$	$MAPE$	0.446 ± 0.055	14 ± 1	129 ± 21
Temperature		$P_m = 0.05$	R	0.881 ± 0.024		
		$P_c = 0.5$	R^2	0.776 ± 0.045		
		$P_{zero} = 0.2$	CE	0.710 ± 0.050		

5.4. Results of the proposed approach

The next step is to use the values of the concepts selected by the SOGA-FCM as inputs to the ANN at the second stage. For the second dataset with one output concept, the Levenberg-Marquardt algorithm was used in learning process. For the first, third and fourth datasets (due to limitation of the Levenberg-Marquardt algorithm implementation in ISEMK) back propagation method with momentum was used. The aim of the experiments is to find the

Table 5: Concepts chosen by the SOGA as the most significant for the analyzed dataset

Dataset	Concepts	n_c	n
1 Electricity	voltage (X_4), global intensity (X_5), energy sub-metering No. 1 (X_6), energy sub-metering No. 2 (X_7) energy sub-metering No. 3 (X_8)	5	8
2 Stock	Istanbul stock exchange national 100 index (X_1), NIK Stock market return index of Japan (X_6), BVSP Stock market return index of Brazil (X_7), EU MSCI European index (X_8)	4	9
3 Bikes	season (X_1), year (X_2), holiday (X_4), weekday (X_5), working day (X_6), weather situation (X_7), temperature (X_8), humidity (X_{10}), wind speed (X_{11}), casual (X_{12}, registered X_{13}, count X_{14})	12	14
4 Temperature	month (X_2), year (X_3), time (X_4), indoor temperatures (X_5, X_6), weather temperature (X_7), relative humidity (X_{10}, X_{11}), lighting (X_{12}), sun light in west facade (X_{17}), sun light in south facade (X_{19}), outdoor temperature (X_{24}), day of the week (X_{26})	13	26

most accurate model for time series prediction. The comparison between the proposed approach (SOGA-FCM-ANN) and previous known methods of prediction (evolutionary FCM and ANN) was presented in Table 6. In fig. 2–5, the
370 predictions of the analyzed approaches and datasets are illustrated.

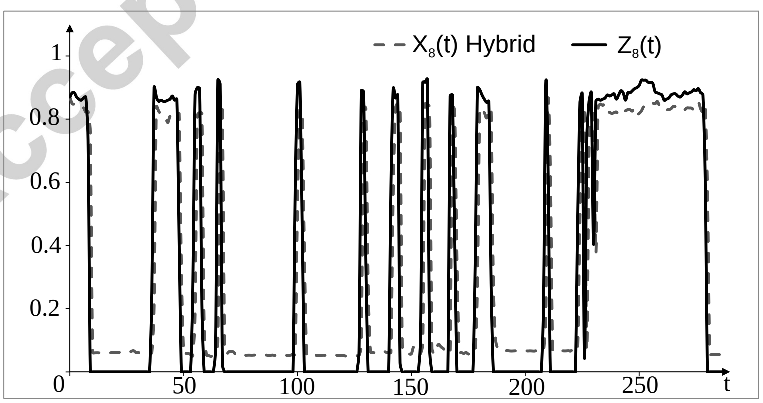


Figure 2: Exemplary results of testing using the first dataset

Table 6: Chosen results of the analysis of evolutionary FCMs (RCGA and SOGA), ANN and Two-stage approach

Dataset	Evolutionary criteria	RCGA	SOGA	ANN	Two-stage
1 Electricity	MSE	**0.021**	**0.021**	0.022	0.022
	$RMSE$	**0.144**	**0.144**	0.149	0.149
	MAE	**0.068**	0.075	0.069	0.069
	$MAPE$	-	-	-	-
	R	**0.617**	**0.617**	0.316	0.315
	R^2	**0.380**	**0.380**	0.100	0.099
	CE	**0.409**	0.393	0.235	0.236
2 Stock	MSE	0.026	0.011	**0.009**	**0.009**
	$RMSE$	0.163	0.105	0.097	**0.095**
	MAE	0.140	0.082	**0.074**	**0.074**
	$MAPE$	0.357	0.207	0.186	**0.183**
	R	0.267	0.223	0.356	**0.395**
	R^2	0.071	0.050	0.130	**0.156**
	CE	-1.560	-0.072	0.090	**0.119**
3 Bikes	MSE	0.032	0.035	0.030	**0.026**
	$RMSE$	0.180	0.188	0.173	**0.162**
	MAE	0.136	0.137	0.128	**0.119**
	$MAPE$	0.645	0.856	0.566	**0.528**
	R	0.650	0.573	0.762	**0.768**
	R^2	0.422	0.328	0.580	**0.590**
	CE	0.354	0.292	0.420	**0.490**
4 Temperature	MSE	0.013	0.009	**0.002**	**0.002**
	$RMSE$	0.113	0.095	0.049	**0.046**
	MAE	0.090	0.077	0.042	**0.039**
	$MAPE$	0.487	0.464	**0.193**	0.196
	R	0.893	0.932	**0.992**	0.991
	R^2	0.797	0.869	**0.983**	0.982
	CE	0.721	0.805	0.948	**0.954**

6. Discussion of Results

In this paper, we conducted a multivariate analysis investigating the advantages and disadvantages of the proposed approach with FCM-based and ANN-based forecasting methods. The most important contribution is on the
375 proposition of the two-stage process, where through the first stage a less complex FCM model is designed after SOGA learning, and next this structure with the most important concepts is incorporated in the ANN structure for learning

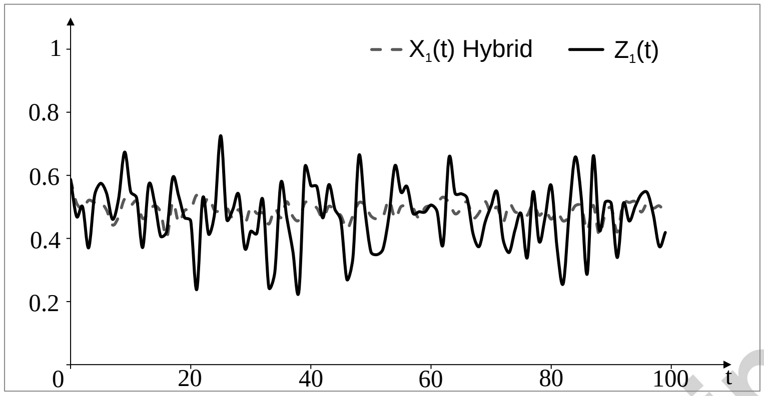


Figure 3: Exemplary results of testing using the second dataset

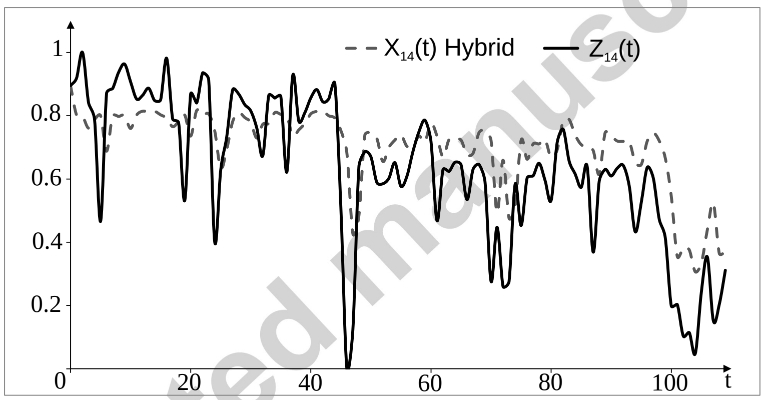


Figure 4: Exemplary results of testing using the third dataset

and testing of time series data.

In table 6, we gathered the results of the analysis of the two efficient evolutionary learning methods for FCMs (RCGA and SOGA), ANN and the proposed two stage model called SOGA-FCM-ANN. We highlighted in bold the best values calculated from the four different prediction methods and the four different case studies. The highlighted values in bold show in each case study the best values achieved for each one of the performance metrics/criteria for prediction. For example, in the third case study on predicting the counts of rental bikes, the highlighted values in bold denote that the two-stage approach has the best values on all seven evaluation criteria (minimum values) compared with the other

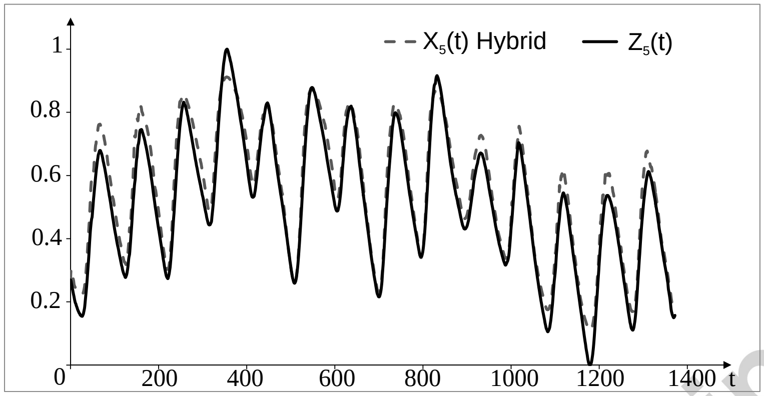


Figure 5: Exemplary results of testing using the fourth dataset

three previously established and known prediction methods.

As shown in Table 6, the two stage model SOGA-FCM-ANN, in most of the cases (three of the four investigated problems), gives better accuracy on prediction with respect to the calculated prediction performance metrics. More specifically, in the third and fourth case study, the proposed approach outperforms for most of the evaluation metrics. We can also observe that the performance of the two-stage process is similar with the performance of the ANN in the first case study, when the initial number of concepts is relatively small (only eight initial concepts were considered). When the number of concepts increases sequentially, it is observed that the proposed SOGA-FCM-ANN is really less complex and performs better in making predictions for all calculated metrics.

Concerning the reduction of complexity, for example, the structure of the FCM model for indoor temperatures prediction learned with the use of RCGA is a complex one as it includes all possible concepts (26 in number). The structure of the same FCM model learned and optimized with the use of SOGA includes the most significant concepts which are 13. This less complex model is really more efficient regarding the prediction of the RCGA-FCM. These selected concepts (13 concepts) with their values defined by the SOGA-FCM were used at the next stage as the inputs for ANN, thus improving the operation of the conventional ANNs in making predictions. Through the conducted experiments with the four different prediction datasets (involving different number of concepts and increased complexity), the results have shown significant outcomes helping us to further investigate the usefulness of two-stage models based on FCMs and ANNs for making multi-variate time series prediction.

Also, through the prediction performance measures, the proposed SOGA-FCM-ANN approach proves to have adequate fitting and forecasting capacity, comparing with the conventional RCGA learning algorithm for FCM and other benchmark ANN models with back propagation or Levenberg-Marquardt learning algorithms. The advantageous characteristic of the selection of the most important concepts defined by SOGA-FCM process helps significantly on increasing the accuracy of prediction finally produced by an ANN. This provides a novelty concerning the FCM theory, as through the SOGA learning a less complex FCM is produced for making predictions, and the concepts derived of this less complex FCM are used as inputs to an ANN structure to make further predictions.

7. Conclusion

In this research study, a two-stage process for multivariate time series prediction was proposed based on evolutionary fuzzy cognitive maps with structure optimization algorithms and ANN-based forecasting methods. The innovation of this research work is the selection of the most important concepts and their interconnections, thus constructing less complex models for time series prediction, with high prediction accuracy comparing with other conventional methods, FCM-based and ANN-based models. The two-stage approach is capable to make better predictions than the conventional evolutionary FCM, by defining the most important concepts through the SOGA-FCM construction approach, which are used at the second stage for improving the operation of the ANN.

To sum up, the proposed approach have an important remark and significant contribution in FCM theory concerning the task of making predictions, where a really large number of factors need to be considered to perform multi-variate analysis. Also, the proposed approach seems very promising for time series forecasting reducing the complexity of the initial network structure in the case of FCM and due to this can be easily implemented in other problems with many variables and increased complexity. Future work is directed to further investigate hybrid models involving other advanced computational intelligent methods for multi-variate time series forecasting where conventional stochastic and statistic modeling approaches are not adequate.

References

[1] S. Ahmadi, N. Forouzideh, S. Alizadeh, and E. I. Papageorgiou, Learning Fuzzy Cognitive Maps using Imperialist Competitive Algorithm, Neural Computing and Applications 26 (6) (2015) 1333–1354.

[2] O. Akbilgic, H. Bozdogan, and M. E. Balaban, A novel Hybrid RBF Neural Networks model as a forecaster. Statistics and Computing. DOI 10.1007/s11222-013-9375-7 (2013).

[3] O. Akbilgic, Hybrid Radial Basis Function Network for selection of variables and forecasting: application to financial securities investment decision (in Turkish), PhD Thesis, Istanbul University (2011).

[4] A. Chaudhuri and K. De, Time Series Forecasting Using Hybrid Neuro-Fuzzy Regression Model, Rough Sets, Fuzzy Sets, Data Mining and Granular Computing 5908 of the series Lecture Notes in Computer Science (2009) 369–381.

[5] H. Fanaee-T and J. Gama, Event labeling combining ensemble detectors and background knowledge, Progress in Artificial Intelligence, Springer Berlin Heidelberg (2013) 1–15.

[6] D. B. Fogel, Evolutionary Computation. Toward a New Philosophy of Machine Intelligence (Third Edition). A John Wiley & Sons, Inc., Publication (2005).

[7] W. Froelich and P. Juszczuk, Predictive Capabilities of Adaptive and Evolutionary Fuzzy Cognitive Maps - A Comparative Study, in: N. T. Nguyen and E. Szczerbicki (Eds.): Intel. Sys. for Know. Management, SCI 252, Springer-Verlag, Heidelberg (2009) 153–174.

[8] W. Froelich and J. Salmeron, Evolutionary learning of fuzzy grey cognitive maps for the forecasting of multivariate, interval-valued time series, International Journal of Approximate Reasoning 55 (2014) 1319–1335.

[9] M. T. Hagan and M. B. Menhaj, Training feedforward Networks with the Marquardt Algorithm, IEEE TRANSACTIONS ON NEURAL NETWORKS 5 (6) (1996) 989–993.

[10] S. Haykin, Neural Networks: A comprehensive Foundation (2nd Edition). Prentice Hall (1999).

[11] G. Hebrail and A. Barard, UCI Machine Learning Repository, http://archive.ics.uci.edu/ml/datasets/Individual+household+electric+power+consumption [8.10.2016].

[12] F. Herrera, M. Lozano, J.L. Verdegay, Tackling Real-Coded Genetic Algorithms: Operators and Tools for Behavioural Analysis, Artificial Intelligence Review 12, (1998) 265–319.

[13] W. Homenda, A. Jastrzebska, and W. Pedrycz, Modeling Time Series with Fuzzy Cognitive Maps, 2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Beijing, China (2014) 2055–2062.

[14] W. Homenda, A. Jastrzębska, and W. Pedrycz, Nodes Selection Criteria for Fuzzy Cognitive Maps Designed to Model Time Series, Advances in Intelligent Systems and Computing 323 (2015) 859–870.

[15] http://epapageorgiou.com/index.php/fcm-software-tools/38-isemk-software-tool [26.04.2016]

[16] A. Jastriebow and K. Poczęta, Analysis of multi-step algorithms for cognitive maps learning, BULLETIN of the POLISH ACADEMY of SCIENCES TECHNICAL SCIENCES 62 (4) (2014) 735–741.

[17] I. Khandelwal, R. Adhikari, and G. Verma, Time Series Forecasting using Hybrid ARIMA and ANN Models based on DWT Decomposition, International Conference on Intelligent Computing, Communication & Convergence (ICCC-2014), Procedia Computer Science 48 (2015) 173–179.

[18] M. Khashei and M. Bijari, A New Hybrid Methodology for Nonlinear Time Series Forecasting, Modelling and Simulation in Engineering, Volume 2011 (2011), Article ID 379121, http://dx.doi.org/10.1155/2011/379121.

[19] M. Khashei, M. Bijari, and G. A. R. Ardali, Improvement of Auto-Regressive Integrated Moving Average models using Fuzzy logic and Artificial Neural Networks (ANNs), Neurocomputing 72 (4–6) (2009) 956–967.

[20] B. Kosko, Fuzzy cognitive maps, International Journal of Man-Machine Studies 24 (1) (1986) 65–75.

[21] W. Lu, W. Pedrycz, X. Liu, J. Yang, and P. Li, The modeling of time series based on fuzzy information granules, Expert Systems with Applications 41 (2014) 3799–3808.

[22] W. Lu, J. Yang, and X. Liu, The Hybrids Algorithm Based on Fuzzy Cognitive Map for Fuzzy Time Series Prediction, Journal of Information & Computational Science 11(2) (2014) 357–366.

[23] J. E. Nash and J. V. Sutcliffe, River flow forecasting through conceptual models part I – A discussion of principles, Journal of Hydrology, 10 (3), (1970) 282–290.

[24] F. Pan, M. Xia, and E. Bai, A new hybrid model for time-series prediction, Cognitive Informatics, ICCI '09. 8th IEEE International Conference on (2009) DOI:10.1109/COGINF.2009.5250728

[25] E. I. Papageorgiou, Fuzzy Cognitive Maps for Applied Sciences and Engineering From Fundamentals to Extensions and Learning Algorithms, Intelligent Systems Reference Library 54, Springer Verlag (2014).

[26] E. I. Papageorgiou, Learning Algorithms for Fuzzy Cognitive Maps - A Review Study. IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and Reviews 42 (2) (2012) 150–163.

[27] E. I. Papageorgiou and W. Froelich, Application of Evolutionary Fuzzy Cognitive Maps for Prediction of Pulmonary Infections, IEEE Transactions on Information Technology in Biomedicine 16 (1) (2011) 143–149.

[28] E. I. Papageorgiou and W. Froelich, Multi-step prediction of pulmonary infection with the use of evolutionary fuzzy cognitive maps, Neurocomputing 92 (2012) 28–35.

[29] E. I. Papageorgiou, K. Poczęta, and C. Laspidou, Application of Fuzzy Cognitive Maps to water demand prediction, Fuzzy Systems (FUZZ-IEEE), 2015 IEEE International Conference on, Istanbul (2015) 1–8.

[30] E.I. Papageorgiou, K. Poczęta, and C. Laspidou, Hybrid Model for Water Demand Prediction based on Fuzzy Cognitive Maps and Artificial Neural Networks, Fuzzy Systems (FUZZ-IEEE), 2016 IEEE International Conference on, Vancouver, Canada, pp. in press.

[31] E. I. Papageorgiou and J. L. Salmeron, A Review of Fuzzy Cognitive Maps Research during the last decade, IEEE Transactions on Fuzzy Systems 21 (1) (2013) 66–79.

[32] G. A. Papakostas, Y. S. Boutalis, D. E. Koulouriotis, and B. G. Mertzios, Fuzzy Cognitive Maps for Pattern Recognition Applications, International Journal of Pattern Recognition and Artificial Intelligence 22 (8) (2008) 1461–1468.

[33] K. Poczęta and A. Yastrebov, Analysis of Fuzzy Cognitive Maps with Multi-Step Learning Algorithms in Valuation of Owner-Occupied Homes,

2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Beijing, China (2014) 1029–1035.

[34] K. Poczęta, A. Yastrebov, and E. I. Papageorgiou, Learning Fuzzy Cognitive Maps using Structure Optimization Genetic Algorithm, 2015 Federated Conference on Computer Science and Information Systems (FedCSIS), Lodz, Poland (2015) 547–554.

[35] J. L. Salmeron, Fuzzy cognitive maps for artificial emotions forecasting, Applied Soft Computing 12 (2012) 3704–3710.

[36] J. L. Salmeron and W. Froelich, Dynamic optimization of fuzzy cognitive maps for time series forecasting, Knowledge-Based Systems 105 (2016) 29–37.

[37] J. L. Salmeron and E. Gutierrez, Fuzzy Grey Cognitive Maps in Reliability Engineering, Applied Soft Computing 12 (12) (2012) 3818–3824.

[38] J. L. Salmeron and E. I. Papageorgiou, A Fuzzy Grey Cognitive Maps-based Decision Support System for Radiotherapy Treatment Planning. Knowledge-Based Systems 30 (1) (2012) 151–160.

[39] H. Song, C. Miao, W. Roel, and Z. Shen, Implementation of fuzzy cognitive maps based on fuzzy neural network and application in prediction of time series, IEEE Transactions on Fuzzy Systems 18 (2) (2010) 233–250.

[40] W. Stach, L. Kurgan, W. Pedrycz, and M. Reformat, Genetic learning of fuzzy cognitive maps, Fuzzy Sets and Systems 153 (3) (2005) 371–401.

[41] W. Stach, W. Pedrycz, and L. A. Kurgan, Learning of fuzzy cognitive maps using density estimate, IEEE Trans. on Systems, Man, and Cybernetics, Part B 42 (3) (2012) 900–912.

[42] G. Słoń, Application of Models of Relational Fuzzy Cognitive Maps for Prediction of Work of Complex Systems, Lecture Notes in Artificial Intelligence LNAI 8467, Springer-Verlag, (2014) 307–318.

[43] G. Słoń and A. Yastrebov, Optimization and Adaptation of Dynamic Models of Fuzzy Relational Cognitive Maps, in: S.O. Kuznetsov et al. (Eds.) RSFDGrC 2011, Lecture Notes in Artificial Intelligence 6743, Springer-Verlag, Heidelberg (2011) 95–102.

[44] A. K. Tsadiras and K. G. Margaritis, An experimental study of the dynamics of the certainty neuron fuzzy cognitive maps, Neurocomputing 24 (1-3) (1999) 95–116.

[45] R. Vidal, J. L. Salmeron, A. Mena, and V. Chulvi, Fuzzy Cognitive Map-based selection of TRIZ trends for eco-innovation of ceramic industry products, Journal of Cleaner Production 107 (2015) 202–214.

[46] L. Wang, H. Zou, J. Su, L. Li, and S. Chaudhry, An ARIMA-ANN Hybrid Model for Time Series Forecasting. Syst. Res. 30 (2013) 244–259.

[47] Y. Xiao, J. Xiao, and S. Wang, A hybrid model for time series forecasting, Human Systems Management 31 (2) (2012) 133–143.

[48] E. Yesil, C. Ozturk, M. F. Dodurka, and A. Sahin, Control Engineering Education Critical Success Factors Modeling via Fuzzy Cognitive Maps, 12th Information Technology Based Higher Education and Training ITHET 2013, Antalya, Turkey (2013).

[49] F. Zamora-Martinez, P. Romeu, P. Botella-Rocamora, and J. Pardo, On-line learning of indoor temperature forecasting models towards energy efficiency, Energy and Buildings 83 (2014) 162–172.

[50] G. P. Zhang, Time series forecasting using a hybrid ARIMA and neural network model, Neurocomputing 50 (2003) 159–175.

[51] X. Zhu, J. Li, D. Wu, H. Wang, and C. Liang, Balancing accuracy, complexity and interpretability in consumer credit decision making: a C-TOPSIS classification approach, Knowledge-Based Systems 52 (2013) 258–267.